



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/663,390	09/15/2003	Dinei A. Florencio	MCS-032-03 (304924.01)	2920
27662	7590	12/26/2007	EXAMINER	
MICROSOFT CORPORATION			LERNER, MARTIN	
C/O LYON & HARR, LLP			ART UNIT	PAPER NUMBER
300 ESPLANADE DRIVE			2626	
SUITE 800			MAIL DATE	
OXNARD, CA 93036			12/26/2007	
			DELIVERY MODE	
			PAPER	

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No.	Applicant(s)	
	10/663,390	FLORENCIO ET AL.	
	Examiner	Art Unit	
	Martin Lerner	2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- Responsive to communication(s) filed on 10 December 2007.
- a) This action is **FINAL**. 2b) This action is non-final.
- Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- Claim(s) 1 to 50 is/are pending in the application.
- a) Of the above claim(s) 22 to 50 is/are withdrawn from consideration.
- Claim(s) _____ is/are allowed.
- Claim(s) 1 to 17, 19, and 21 is/are rejected.
- Claim(s) 18 and 20 is/are objected to.
- Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- The specification is objected to by the Examiner.
- The drawing(s) filed on _____ is/are: a) accepted or b) objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
a) All b) Some * c) None of:
 - Certified copies of the priority documents have been received.
 - Certified copies of the priority documents have been received in Application No. _____.
 - Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- Notice of References Cited (PTO-892)
- Notice of Draftsperson's Patent Drawing Review (PTO-948)
- Information Disclosure Statement(s) (PTO/SB/08)
Paper No(s)/Mail Date _____.
- Interview Summary (PTO-413)
Paper No(s)/Mail Date. _____.
- Notice of Informal Patent Application
- Other: _____.

DETAILED ACTION

Election/Restrictions

1. Applicants' election of Group I, Claims 1 to 21, in the reply filed on 10 December 2007 is acknowledged. Because Applicants did not distinctly and specifically point out the supposed errors in the restriction requirement, the election has been treated as an election without traverse (MPEP § 818.03(a)).
2. Claims 22 to 50 are withdrawn from further consideration pursuant to 37 CFR 1.142(b) as being drawn to a nonelected invention, there being no allowable generic or linking claim. Election was made **without** traverse in the reply filed on 10 December 2007.

Claim Rejections - 35 USC § 102

3. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

4. Claims 1, 7 to 11, and 21 are rejected under 35 U.S.C. 102(b) as being anticipated by *Hardwick et al.*

Regarding independent claim 1, *Hardwick et al.* discloses a system for playing back an audio or speech signal (column 4, lines 64 to 67), comprising:

“storing received audio data to a signal buffer” – A/D converter 502 will provide 200 samples to buffer 506 every $\frac{1}{4}$ second (column 9, lines 13 to 15; column 9, lines 29 to 33: Figure 5A);

“outputting parts of the signal present in the signal buffer as needed for signal playback” – modulator 508 will take a packet of 200 samples from TSM 512 (and thus from buffer 506) every $\frac{1}{4}$ second (column 9, lines 29 to 33: Figure 5A);

“analyzing the contents of the signal buffer” – the apparatus solves the data rate mismatch problem by comparing the fullness of the buffer over time to two limit levels, a high and a low level, are indicated by lines 602 and 604 respectively (column 9, lines 56 to 62: Figures 5A, 6A, and 6B);

“stretching at least part of the signal present in the signal buffer when the analysis of the contents of signal buffer indicates that the length of the signal in the signal buffer is less than a predetermined threshold” – if buffer 506 becomes too empty, TSM (time scale modification) module 512 starts to take fewer than N samples from the buffer and transforms N-y samples into N samples until the level of samples in buffer 506 returns to a level between the low and high levels (column 10, lines 40 to 53: Figures 5A, 6A, and 6B);

“compressing at least part of the signal present in the signal buffer when the analysis of the contents of the signal buffer indicates that the length of the signal in the signal buffer is greater than a predetermined threshold” – if buffer 506 becomes more full than the high limit, this fact is communicated to TSM (time scale modification) module 512 over control line 514 and TSM module 512 will take more than 200

samples, e.g. 206 samples, from the buffer 506 for every 2000 clock pulses of clock 510, and transforms this extra large packet size into a packet of 200 samples so as to compress N+x samples into N samples (column 9, line 62 to column 10, line 32: Figures 5A, 6A, and 6B).

Regarding independent claim 8, *Hardwick et al.* discloses a system for playing back an audio or speech signal (column 4, lines 64 to 67), further comprising:

"receiving and decoding data frames of an audio signal transmitted across a packet-based network" – the invention can also be applied to an apparatus that includes a demodulator 558, a buffer 556, and a TSM module 562 (column 11, lines 31 to 37: Figure 5B); speech can be coded by modeling speech according to certain parameters with "Improved Multi-Band Excitation" ("IMBE"), where each window of digitized signals is transformed by a Fast Fourier transform to obtain three sets of parameters (fundamental frequency, spectral envelope, and voiced character) (column 7, line 66 to column 9, line 6); samples are transmitted as packets (column 9, lines 14 to 18), so the channel is "a packet-based network"; implicitly, if speech is coded by IMBE, then demodulator 558 provides decoding of speech before playback;

"outputting one or more of the decoded frames present in the signal buffer when the analysis of the contents of the signal buffer indicates that the length of the signal in the buffer is between a predetermined minimum and a predetermined maximum buffer size" – the apparatus solves the data rate mismatch problem by comparing the fullness of the buffer over time to two limit levels, a high and a low level, are indicated by lines

602 and 604 respectively (column 9, lines 56 to 62: Figures 5A, 6A, and 6B); if there is no mismatch, the signal passes along without any time scale modification, which is accomplished by routing the signal around TSM 322 (column 12, lines 40 to 45: Figures 5B and 7: Step 708).

Regarding claims 7 and 21, *Hardwick et al.* discloses an objective is to compensate for clock drift (column 9, lines 29 to 45: Figure 5A).

Regarding claim 9, *Hardwick et al.* discloses that samples move through buffer 556 from demodulator 558 to TSM 562 and D/A converter 552 (column 11, lines 31 to 44: Figure 5B); thus, any samples output from buffer are removed from the buffer.

Regarding claim 10, *Hardwick et al.* discloses that the apparatus is designed to compensate for portions of an utterance that are lost because the channel is not prepared to accept it when it is delivered or because the channel becomes incapable of receiving a signal due to a dropout situation (column 7, lines 7 to 60); thus, portions of utterances that are lost because the channel is late in opening up are equivalent to “late loss packets”, and an apparatus that prevents these situations performs “packet loss concealment”, implicitly.

Regarding claim 11, *Hardwick et al.* discloses stretching a packet by transforming N-y samples into N samples (column 10, lines 40 to 53: Figures 5A and 5B); generally, “jitter” is defined as “an unwanted variation of signal characteristics”, where “clock jitter” is one known form of “jitter”; thus, by compensating for clock drift to control how

samples are stretched or compressed from a buffer (column 9, lines 29 to 67), *Hardwick et al.* provides "automatic jitter control as a function of buffer content."

Claim Rejections - 35 USC § 103

5. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

6. Claims 2 to 4, 12 to 15, and 19 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Hardwick et al.* in view of *Chong-White et al.*

Concerning claims 2, 12, and 13, *Hardwick et al.* notes that frames of speech may be analyzed to determine voiced character in IMBE (column 8, line 50 to column 9, line 6), and it is known that voicing character involves periodic content for voiced vowels and aperiodic content for unvoiced consonants. *Hardwick et al.* does not expressly provide for analyzing a contents of the signal buffer from a group including periodic content and aperiodic content, or voiced and unvoiced frames, prior to stretching decoded frames. However, *Chong-White et al.* teaches enhancing speech intelligibility using variable-rate time-scale modification, where vowel sounds (often referenced as voiced speech) and consonant sounds (often referenced as unvoiced speech) are processed from a buffer 702 so that some segments have lengthened time durations, corresponding to stretching, and other segments have compressed time durations, corresponding to compression. (Column 3, Lines 7 to 10; Figures 1 and 7) Specifically,

formant transitions are emphasized through time expansion, and vowel segments are compressed. (Column 7, Lines 48 to 65: Figures 7 and 8) The objective is to enhance speech intelligibility due to consonant confusions in the presence of bandwidth reduction and packet loss. (Column 1, Line 60 to Column 2, Line 20) It would have been obvious to one having ordinary skill in the art to analyze contents of a signal buffer including periodic, aperiodic, voiced, and unvoiced segments prior to stretching or compressing as taught by *Chong-White et al.* in an apparatus for maintaining data rate integrity of *Hardwick et al.* for a purpose of enhancing speech intelligibility in the presence of bandwidth reduction and packet loss.

Concerning claims 3 and 14, *Chong-White et al.* teaches stretching of segments involves searching using a cross-correlation to find a segment within a given tolerance ("identifying at least one segment . . . as a template") that has a maximum similarity ("exceeds a predetermined threshold") to the continuation of a last extracted segment (column 8, lines 39 to 43); the segment is matched with another segment using cross-correlation and waveform similarity criterion, and the segment and the best-matched segment are blended together by overlapping and adding the two segments together ("aligning and merging") (column 9, lines 18 to 41: Figure 8).

Concerning claims 4 and 15, *Chong-White et al.* teaches stretching consonants and unvoiced fricatives (column 7, lines 38 to 55), which are segments having "unvoiced" or "aperiodic" content, to increase speech intelligibility.

Concerning claim 19, *Chong-White et al.* teaches compressing a vowel following a consonant (column 7, lines 54 to 56), where a vowel is a "voiced frame"; procedures

for stretching and compressing both involve searching using cross-correlation to find a segment having maximum similarity, and blending the best-matched segment together by overlapping and adding (column 7, lines 39 to 43; column 8, lines 18 to 41; Figure 8); one skilled in the art would know that the same procedure could be applied to "cutting out" matching signals for compressing and "inserting" matching signals for stretching.

7. Claims 5 to 6 and 16 to 17 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Hardwick et al.* in view of *Chong-White et al.* as applied to claims 1, 2, 4, 8, and 15 above, and further in view of *Unno et al.*

Hardwick et al. discloses processing speech into a frequency domain using a Fast Fourier transform in speech modeling by "Improved Multi-Band Excitation" ("IMBE") (column 8, lines 2 to 49), which implicitly involves decoding the speech into a time domain by an inverse Fast Fourier transform. Moreover, *Chong-White et al.* teaches application to a coder operating by Mixed Excitation Linear Prediction (MELP) (column 2, Lines 21 to 25). *Hardwick et al.* omits introducing a random rotation of the phase into frequency domain signals by applying at least one LPC filter to compute an LPC residual, computing at least one FFT from the LPC residual, introducing a random phase rotation into the coefficients, computing an inverse FFT, and applying an inverse LPC filter to the LPC residual to create at least one synthetic segment. However, *Unno et al.* teaches an enhancement to a mixed excitation linear predictive (MELP) coder, where one embodiment involves taking the Fourier magnitude of an LPC residual 23, introducing a random phase 64, performing an inverse DFT 93, and producing a mixed

excitation signal 95 (column 11, lines 53 to 66: Figure 9). One skilled in the art would know that an LPC residual is then processed through an LPC synthesis filter to create synthesized speech. (Figures 2C and 8) An objective is to enhance the coded speech quality of a MELP coder for plosives (column 2, lines 31 to 56), which are unvoiced or aperiodic speech. It would have been obvious to one having ordinary skill in the art to perform the technique of random phase rotation of a frequency domain LPC residual as taught by *Unno et al.* in an apparatus for maintaining data rate integrity of *Hardwick et al.* for a purpose of enhancing the coded speech quality of plosives in an coder operating in accordance with MELP.

Allowable Subject Matter

8. Claims 18 and 20 are objected to as being dependent upon a rejected base claim, but would be allowable if rewritten in independent form including all of the limitations of the base claim and any intervening claims.

Conclusion

9. The prior art made of record and not relied upon is considered pertinent to Applicants' disclosure.

Leitch et al., Coorman et al., Goldhor et al., Vargo et al., Christensen et al., Overby, Jr. et al., and Verreault disclose related art.

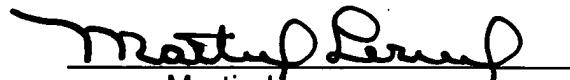
Any inquiry concerning this communication or earlier communications from the examiner should be directed to Martin Lerner whose telephone number is (571) 272-

7608. The examiner can normally be reached on 8:30 AM to 6:00 PM Monday to Thursday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, David R. Hudspeth can be reached on (571) 272-7843. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

ML
12/20/07



Martin Lerner
Examiner
Group Art Unit 2626